Sampling and Spatial Representation

The Acquisition and Content of GIS Data

Karl Hermann, EPA/Region 8 - *Presenter*

Spatial Representation

Content

- How spatial data is handled in a GIS
- The importance of scale
- The temporal dimension
- Portrayal of thematic layers

Acquisition

- Primary techniques
- Associated issues

Goal of data representation in GIS:

To provide spatial characterization of thematic 'layers' at desired scales and level of detail

Usually, the representation is an abstraction of reality

- a point for a city
- a line for a highway

Three Dimensions of the World:

- Spatial - variation in place
- Temporal - variation in time
- Thematic - variation in a characteristic

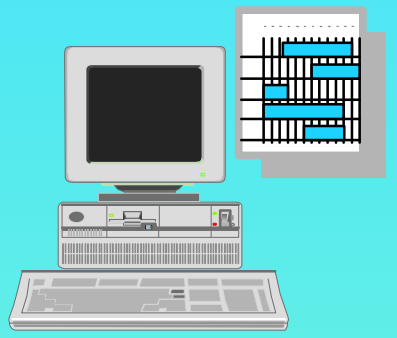
- All GIS handle spatial and thematic
- Many handle temporal

Spatial Dimension

"How to Represent Spatial Features in a GIS"

Geographic features are usually represented in a GIS as spatial features with x,y, and sometimes, z and t (time) coordinates.

"Digital World"



x,y,z

Vector Data

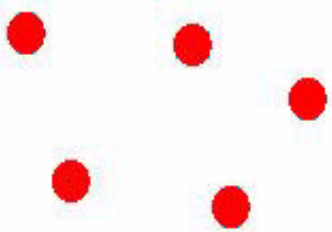
- points
- lines (arcs)
- areas (polygons)

Raster Data

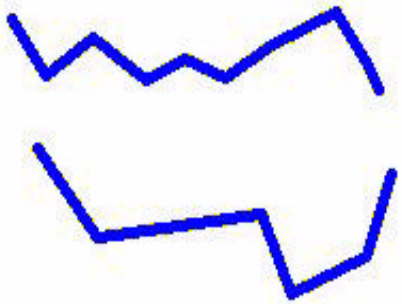
- grids (grid cells or pixels)

Geographic Information Systems

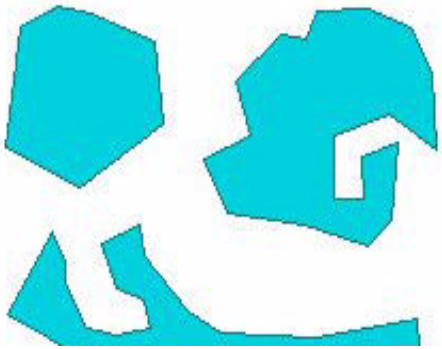
Spatial Features



points



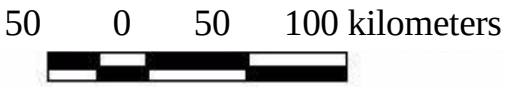
lines



polygons

Attributes

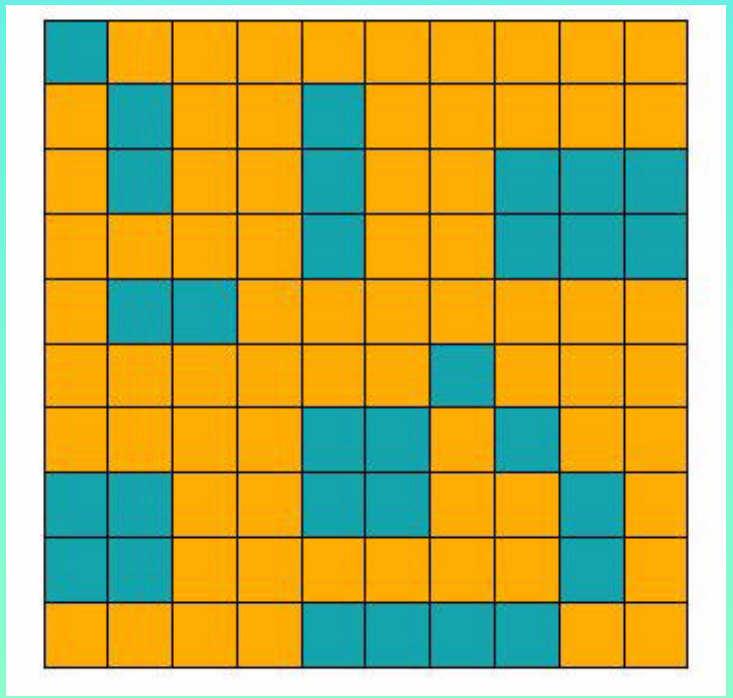
Counties	Fips	St	Cntyname
0	0		
2826	30053	MT	LINCOLN
3056	38067	ND	PEMBINA
2802	30029	MT	FLATHEAD
3008	38019	ND	CAVALIER
2808	30035	MT	GLACIER
3084	38095	ND	TOWNER
3068	38079	ND	ROLETTE
2874	30101	MT	TOOLE
2998	38009	ND	BOTTINEAU
2824	30051	MT	LIBERTY
2814	30041	MT	HILL



Raster Data

Fixed cell size

grid cells or pixels



Coordinate References

true shape of the Earth (topography and geoid)
ellipsoid (spheroid)
datums (horizontal and vertical)
projection
coordinate system

Ellipsoids
- Clark 1866
- GRS 1980

Datums
- NAD27
- NAD83

Projections and Coordinate Systems

- definition with respect to latitude and longitude
- truth on given latitude(s) or longitude(s)

Projections

Albers

UTM

State Plane

etc.

Coordinate Systems

units (meters, feet)

origin

false eastings or northings

Scale

"How do we want to view the data?"

Level of Abstraction / Detail

Intended Use

Scale is a Representative Fraction

smaller scale vs. larger scale

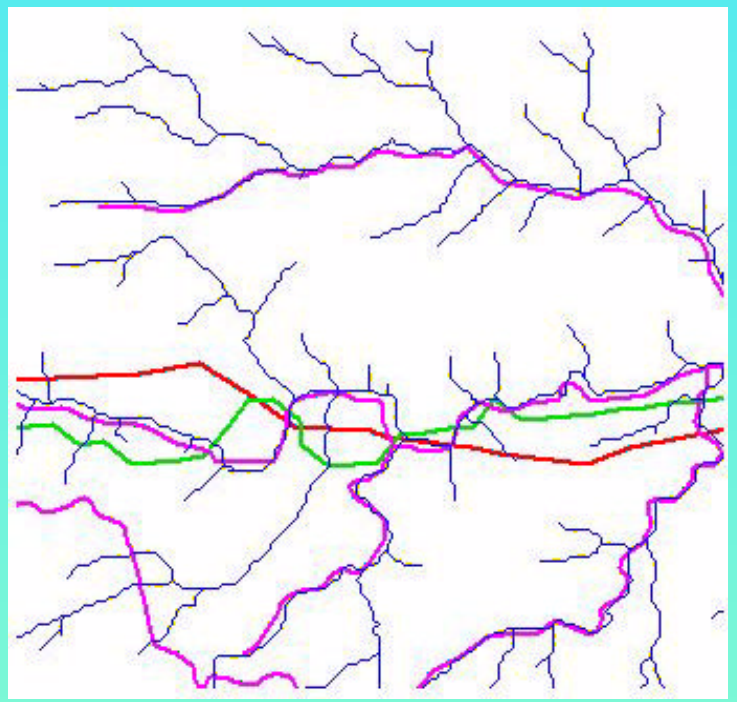
1/2,000,000 < 1/100,000

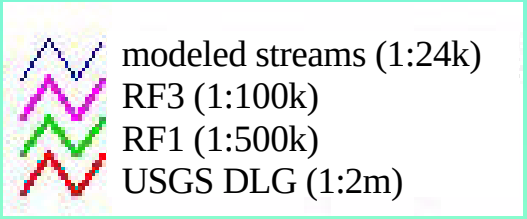
smaller (coarser) scale data is NOT bad

GIS Data Quality Objectives:

Locational accuracy DQO varies with scale

Hydrography Example: Four Scales

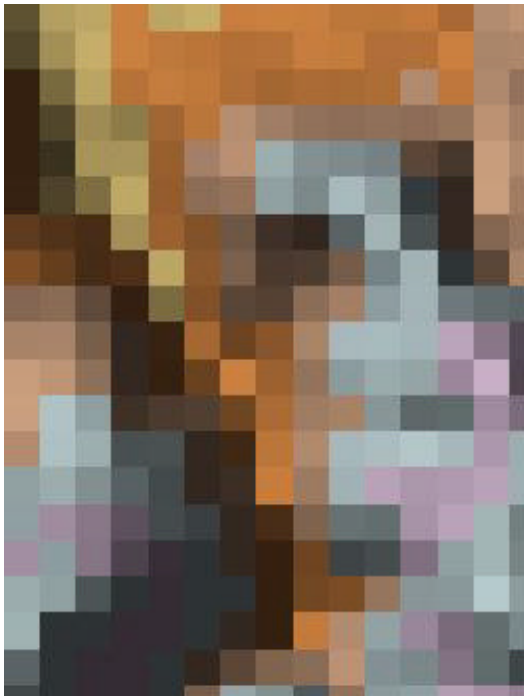




Scale

Digital Elevation Model Example shown at three scales

900 meter cell size
~ 1:2000000 source

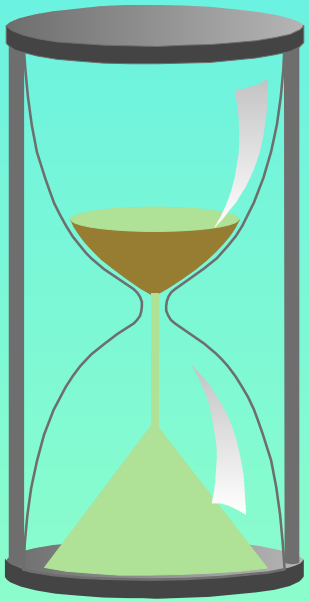


90 meter cell size
~ 1:250000 source

30 meter cell size
~ 1:24000 source

Temporal Dimension

"How to Represent Time in a GIS"



Handling Time and Date

Metadata record of the theme

Attribute (for fixed or dynamic themes)

GIS with temporal handling ability (update history)

GIS in the future is x,y,z,t

Temporal Issues

Date of source

Date of labeling / attribution

Dynamic nature of theme

Thematic Dimension

"How to Represent Spatial Themes in a GIS"

Aggregation is *f*(scale)

smaller scale

aggregated units

less detail

Precise Boundaries
vs.
Ecotones

Degree of "fuzziness"

discrete classes
- *f*(fence lines)
- *f*(ownership lines)
- artifical features

continous variation
- elevation
- soils
- temperature

Classification Issues

Mixed types per unit (area or pixel)

Normalization (account for unequal areas)

Definitions (density, width, priority)

Count
Identity
Order
Interval

GIS Data Acquisition

"How do we get Data into a GIS?"

GIS Data Acquisition

Spatial data collection methods

- Direct Measurement
- Remote Sensing
- Secondary Measurements

Mapping and sampling techniques are included in all three methods

Method Variables

- Accuracy and Precision
- Scale of Materials and Analysis
- Resolution of Materials/Instruments
- Orthographic Integrity
- Interpretation Skills
- Date and Time

GIS Data Acquisition Techniques

Mapping

- Surveying, remote sensing, ...
- Interpretation and delineation
- Digital conversion

Sampling

- Thematic assignment
- Interpolation / extrapolation
- "Training data" for mapping

Mapping

- Compilation, interpretation, and/or delineation
- Field data, remotely-sensed data, and/or other
- Need to use or define proper geographic base

- Mapping standards
- Interpretation/classification SOPs
- Quality assurance

Sampling Purposes

- Capturing continuous variation
- Basis for population estimates
- "Representative" detail

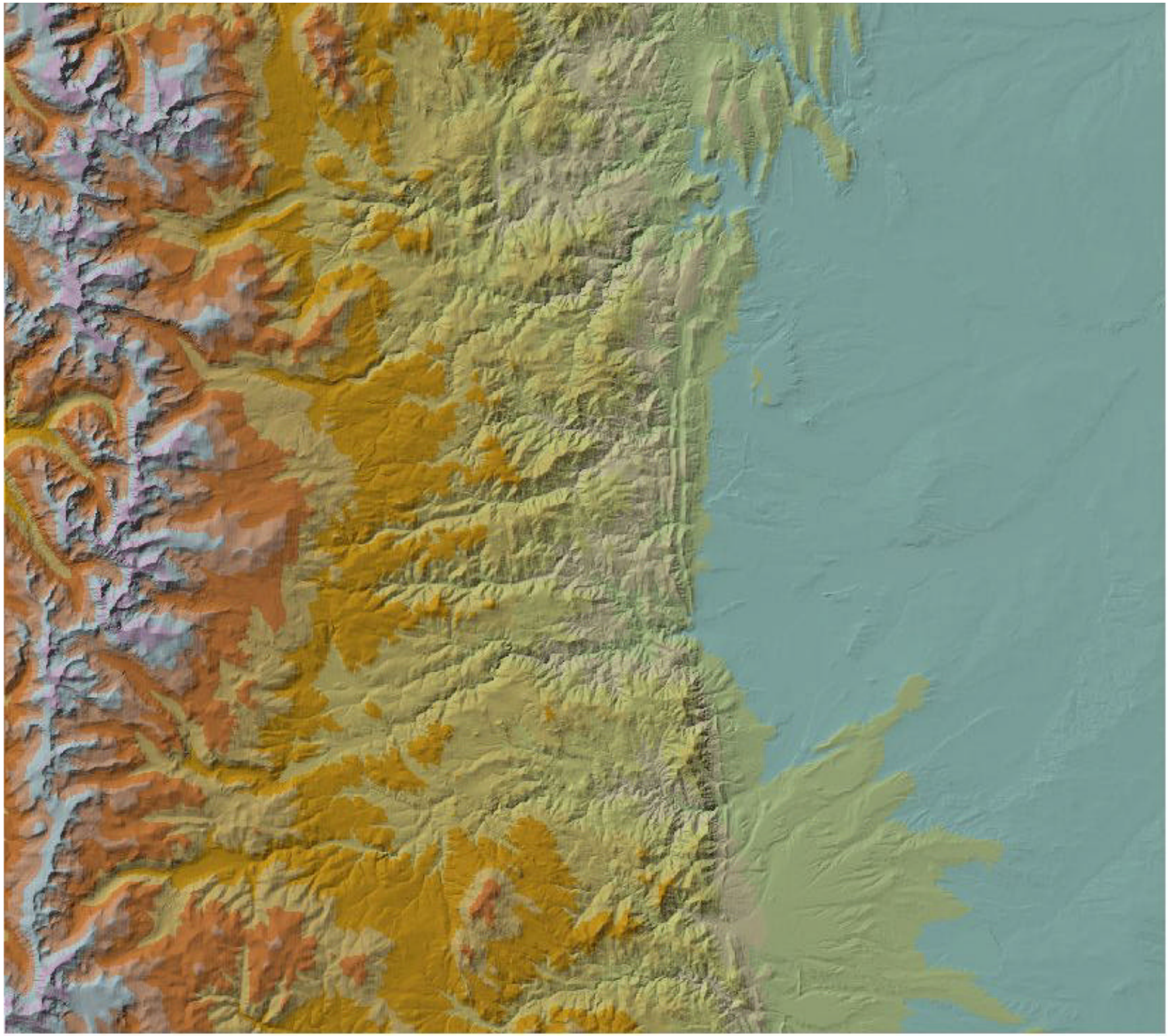
Sampling

A Method of Capturing Continuous Variation

- Examples are: elevation, atmospheric temperature and pressure, natural vegetation and soil type
- Variation can be captured through a sampling design

Measurements in designs can include:

- taking measurements at sample points
- taking measurements along a transect
- characterizing an area frame



Direct Measurements

Locational

surveying
GPS
feature identification
address identification

Attributes

Qualitative

windshield survey
observation by design

Quantitative

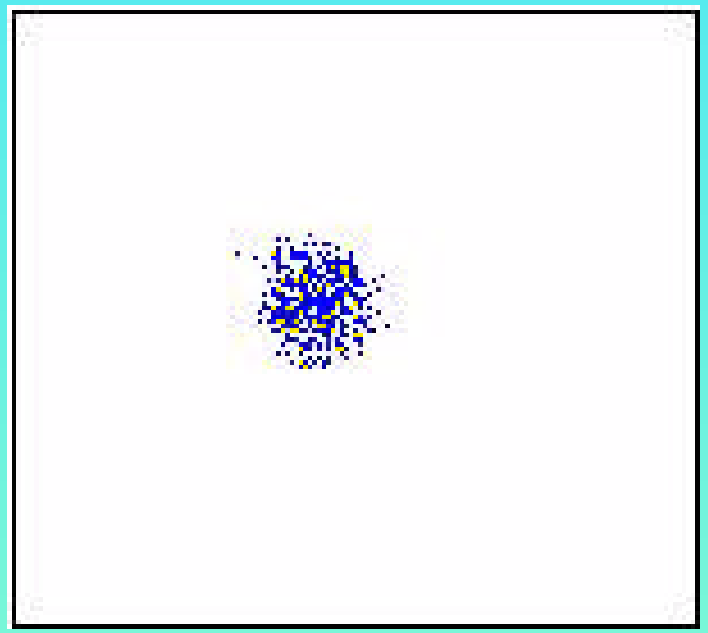
sample design protocol

Global Positioning System

Locational estimation

Sample-based procedure

GPS Sample Data



Raw - Uncorrected

Corrected

GPS Procedures
SOP and QA Considerations (Plan)

standard parameters
checklist
adequate sample size

differential correction
sample mean
"reality check"
GPS unit verification

GPS Quality Assurance

Once a year, for each unit:

- Procedure applied to known geodetic control
- Multiple units, multiple times, multiple controls
- Adequate samples for analysis of results

Accuracy assessment

Precision assessment

Repeatability assessment

Remote Sensing

sensor resolution

geometric correction / registration

interpretation / classification

aerial photography

digital scanner-based

digital satellite-based

Remote Sensing Issues

Analog

- Aerial photo interpretation skills
- Repeatability

Digital

- Classification confusion

- Geometric rectification
- Mapping unit size
- Resolution
- Accuracy assessment

Photogrammetric Mapping

Stereo model (3-D viewing)

Registration to geodetic control

Interpretations / delineations performed from model

Secondary Measurements

Address matching

Map interpolation

Feature matching / transfer

Orthographic integrity

Geodetic control

Secondary Measurement Issues

Lack of SOP

Poor control base

Unknown / uncertain accuracy

Incorrect interpolation / interpretation

variability (inconsistency)

repeatability

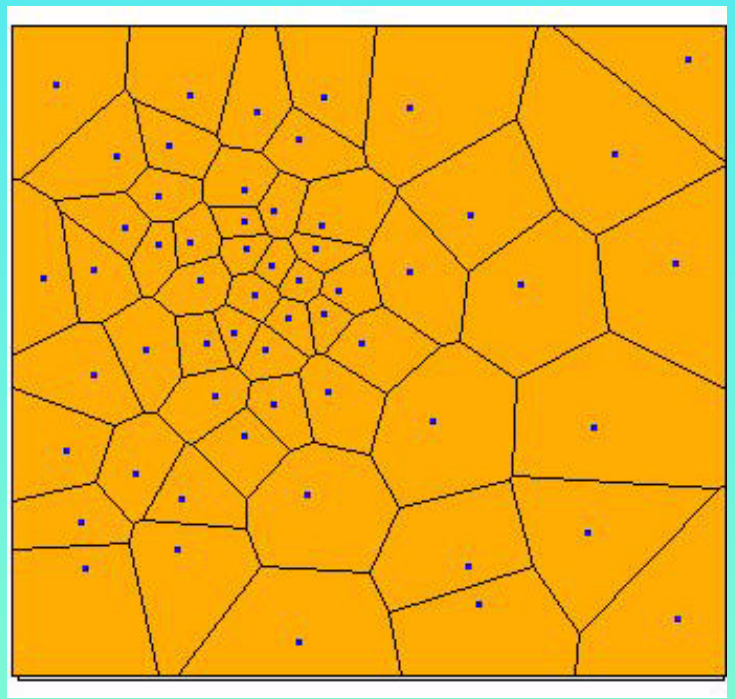
compounded error

Interpolation and Extrapolation

Between and beyond samples
Sample design and/or point distribution
is critical to success

Point source input, modeling method (software version) and metadata are very important.

Area Assignment Based on Samples



Thiessen Polygons need adequate density and distribution

Conclusions

- GIS data content and issues
- How GIS data represent "real world" features
- Importance and meaning of scale
- Mapping and sampling techniques and issues

Questions / Comments